and half-open TCP sessions.



## What is claimed is:

1	1. A method for use with a stateful packet processing device of a computer
2	network for mitigating effects of a network overload against said device, said method
3	operable to free memory used to store information about communications sessions
4	managed by said device, said method comprising the steps of:
5	classifying session cache entries made in memory into different cache classes,
6	according to one or more characteristics of those entries;
7	determining when said device is under network overload;
8	selecting session cache entries for deletion and deleting them thereby freeing
9	associated memory when said device is under network overload;
0	determining when sufficient memory has been freed, such that said cache entries
1	are no longer deleted.
1	2. The method of Claim 1, wherein said characteristics for said step of
2	classifying are selected from the group consisting of: whether the session is dropped by
3	the device, whether the session is audited by the device, IP protocol of the session, ICMP
4	type and code used in the session, TCP ports used in the session, UDP ports used in the
5	session, and whether the session is a half-open TCP session.
1	3. The method of Claim 1, wherein certain of said characteristics of the
2	session may be identified as "any", wherein any session matches a particular criterion for
3	classification.
1	4. The method of Claim 1, wherein predefined cache classes are selected
2	from the group consisting of:

dropped and unaudited sessions, dropped and audited sessions, ICMP sessions,

3

4

5

6

7

8

9

10

11

12

13

14

1

2

3

deletion.

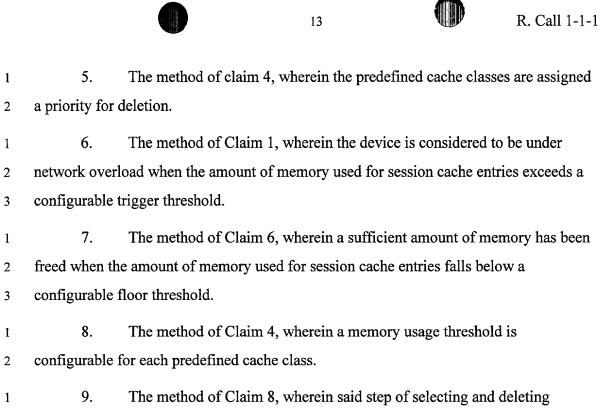
threshold.

10.

includes the steps of:

usage threshold;

entries for each cache class;



retrieving from a database the amount of memory used to store session cache

ordering each cache class according to its deletion priority;

wherein, otherwise, all entries of said given class are selected for deletion; and

recognizing each cache class whose memory usage exceeds an associated memory

selecting for deletion according to said ordering step some fraction of entries of a

given cache class if said deletion brings said total cache memory usage below said floor,

continuing said step of selecting for deletion until it is determined that either

deleting all the entries selected for deletion would bring the total cache memory usage

cache classes whose memory usage does not exceed said associated memory usage

below the floor threshold, or all entries in all defined cache classes have been selected for

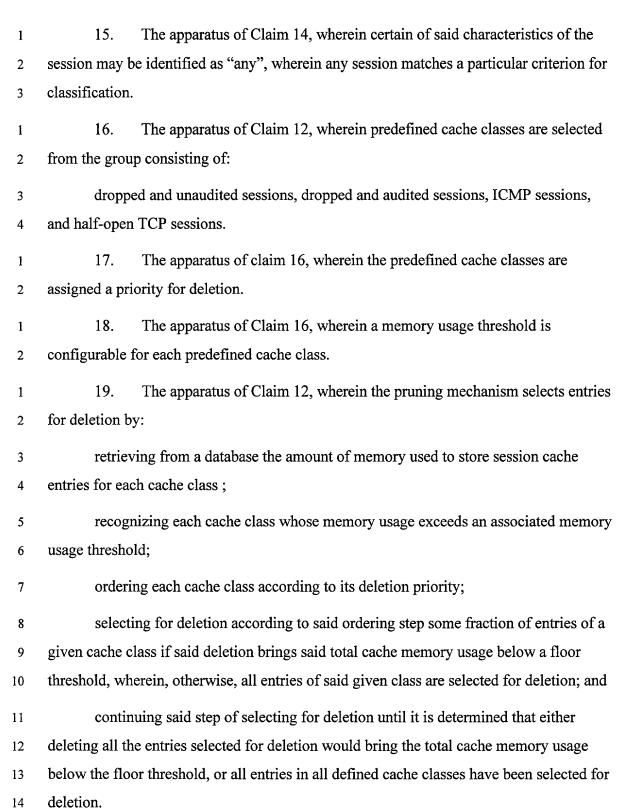
The method of Claim 9, wherein said step of ordering includes ordering

5

1	11. The method of Claim 9, wherein configuration data for the thresholds may
2	be supplied in a normalized fashion and be adaptively applied to the device, depending on
3	the amount of memory on the device.
1	12. An apparatus for use with a stateful packet processing device of a
2	computer network for mitigating effects of a network overload against said device, said
3	apparatus operable to free memory used to store information about communications
4	sessions managed by said device, said system comprising:
5	a classification component operable to determine, for each session cache entry, the
6	cache class to which that entry belongs according to one or more characteristics of the
7	entry;
8	a memory management database for tracking the amounts of memory used for
9	each category of entry, as well as for tracking the total amount of memory used for all
10	entries;
11	a pruning component that is used to select and delete entries; and
12	a processor for determining when said device is experiencing network overload
13	and selecting specific cache session entries for deletion until sufficient memory has been
14	freed.
1	13. The apparatus of Claim 12, wherein information kept in the memory
2	management database is updated each time a new cache entry is created or deleted by the
3	device.
1	14. The apparatus of Claim 12, wherein said characteristics for said step of
2	classifying are selected from the group consisting of: whether the session is dropped by
3	the device, whether the session is audited by the device, IP protocol of the session, ICMP

type and code used in the session, TCP ports used in the session, UDP ports used in the

session, and whether the session is a half-open TCP session.



- 20. The apparatus of Claim 19, wherein said step of ordering includes cache classes whose memory usage does not exceed said associated memory usage threshold.
- 21. The apparatus of Claim 19, wherein the pruning mechanism operates by making only one pass through a list of session cache entries in said device.
- 22. The apparatus of Claim 12, wherein a trigger threshold and floor threshold corresponding to said total memory usage are adjustably configurable.
- 23. The system of Claim 12, wherein the memory usage statistics are collected using the Simple Network Management Protocol (SNMP).
- 24. The apparatus of Claim 12, wherein the pruning mechanism, when it has to delete some fraction of the entries in a given cache class, approximates the fraction b/t (where b is the total number of bytes of memory that must be freed and t is the total number of bytes of memory used to hold session cache entries for that cache class) with another fraction p/q, where  $p \ge 1$  and q is likely to be small relative to the total number of cache entries in that class; and then frees p entries out of every q entries in that cache class on the list of session cache entries.

memory threshold value is reached.

25. A cache management system used in connection with session-type packet
processing devices of a computer network, said system comprising:
a memory management database for storing communication traffic classification
and memory threshold values;
a memory monitor for tracking overall memory usage and determining when said
memory threshold values stored in said memory management database are reached;
a cache classifier used to determine a class into which a given session of
communications traffic falls; and
a pruner mechanism for selecting and pruning selected sessions of said packet
processing device in accordance with said communication traffic classification and

26. The system of Claim 25 wherein said prune selector is operable to selectively prune sessions of an ordered overlimit class if the memory used by said class is greater than the difference between a global ceiling threshold and a global floor threshold.

memory thresholds programmed in said memory management database when said

- 27. The system of Claim 26, wherein said prune selector is operable to prune all sessions of said overlimit class if the memory used by said class is less than the difference between said global ceiling threshold and said global floor threshold.
- 28. The system of Claim 27, wherein a next highest priority class is examined to determine if memory used by said class is greater than a remaining difference between said global ceiling threshold and said global floor threshold, said next highest priority class being selectively pruned if said difference is greater than said remaining difference.
- 29. The system of Claim 28, wherein said prune selector is operable to prune all sessions of said next highest priority class if the memory used by said class is less than said remaining difference.
- 30. The system of Claim 25, wherein said devices are selected form the group consisting of: network firewalls, routers, switches and hosts.